provide reliable coefficient estimates. Data obtained from real world observations most often contain some random influences, and, in regression analysis, we seldom have the luxury of drawing repeated samples to isolate the signal from the noise in the data. We need to work with real world observations, with all of their peculiarities. When a modeler adds and removes variables to improve the fit to a specific set of data, the peculiarities of the single sample can have an exaggerated impact on the result. The result may be a better "fit" to the specific sample, but there is less confidence that the model provides a good fit outside of the sample predictions. In regression analysis, the data is meant as a test and quantification of hypothesized relationships, not to fit an equation to a specific sample. It is unclear how and where the variation in the dependent variable caused by a deleted variable is attributed to other coefficients. By eliminating "local DEMs" from the model, the FCC further limits that reliability of its models.

Finally, the FCC proposes averaging the results of two different model specifications for toll DEMs. This is inappropriate and serves to demonstrate that the FCC does not have confidence in either of its model specifications.

C. Highly correlated data

After considering the model specification and the key drivers of the dependent variable, the next step in carrying the analysis forward is finding data that are suitable to represent these drivers. The variables considered for the FCC's proposal are switched lines, local dial equipment minutes (DEMs), special access channel equivalents, and toll DEMs.

According to the FCC, the variable for local DEMs was removed because it is highly correlated with switched lines. In the words of the FCC, "the number of switched lines and local DEMs were so highly correlated that it did not increase the explanatory power of the model to include both variables."[17] This is a problem caused by the data, and it is not restricted to the high correlation between switched lines and local usage. Table 10

[17] NPRM at ¶ 219, p. 87.

shows the pair-wise correlation between the variables used to create the FCC's explanatory variables.

Table 10

Correlation Between Variables Used by the FCC to Create Explanatory Variables

	Switched lines	Special lines	Local DEMS	Toll DEMS
Switched lines	1.000			
Special lines	0.901	1.000		
Local DEMS	0.977	0.870	1.000	
Toll DEMS	0.977	0.906	0.922	1.000

A value of 1.000 indicates perfect correlation between two variables, as can be seen in the correlation between each variable and itself. A correlation that is close to one indicates that the variations in two variables are very similar – i.e., variations in one variable are closely matched in directions and relative degree by variations in the other. Table 10 shows very high correlations among any two of the explanatory variables considered by the FCC, including switched lines, special lines, local DEMS and toll DEMS. This means that the variation in any variable that is different from the variations in other variables is very small. This creates the problem that the FCC observed. After the initial explanatory variable is selected, each additional variable contributes little to the explanatory power of the model. This does not mean that variables should be deleted, but it does mean that it is very difficult to separate the causative effects of different variables on expense accounts, and this is a problem that renders the FCC's equations inadequate for the purpose at hand. Peter Kennedy explains the implications of highly correlated explanatory variables as follows:

> When the regressors [explanatory variables] are highly correlated, most of their variation is common to both variables, leaving little variation unique

> to each variable. This means that the OLS procedure has little information to use in making its coefficient estimates, just as though it had a very small sample in which the independent variable did not vary much. Any estimate based on little information cannot be held with much confidence...It is this uncertainty as to which variable deserves the credit for the jointly explained variation in the dependent variable that creates the uncertainty as to the true values of the coefficients being estimated.[18]

With highly correlated explanatory variables, the overall fit of the model can be high and the individual coefficients can even appear significant, but, nonetheless, it is impossible to have confidence in the accuracy of the coefficient estimates. Because the FCC's proposed cost attribution is based on coefficient estimates that are derived from highly correlated data, it is not possible to place the necessary confidence in the accuracy of the results.[19]

Examination of the confidence intervals around the FCC's forecasts reveals the extent of the multicollinearity problem. The first entry of Table 11 shows (for FCC specification 1) that the FCC's unadjusted estimate of the Services expense associated with the local switched access line equals $17.28. Columns 2 and 3 indicate that with 95 percent confidence the true mean of Services expense per switched access line falls between $6.02 and $28.54, a range of plus or minus 65%. The confidence interval for expense associated with special access lines is even wider, extending from -$15.38 to $43.51, a range of plus or minus 209%. The confidence interval for toll DEMS is also substantial. Even though it is not possible to have confidence in the precision of the individual coefficient estimates, the forecast for the overall service expense is reasonably reliable, with an error of within plus or minus 5%. These results demonstrate that the FCC's methodology provides a reasonable forecast for the overall amount of service expense, but it does not provide meaningful estimates of the individual causal relationships between the individual independent variables and service expense. Results for the other accounts follow a similar pattern and are provided in Attachment II.

[18] Kennedy, Peter, *A Guide to Econometrics*. The MIT Press. Cambridge, Massachusetts. 1992. P. 178.

Table 11

Forecast Confidence Intervals for Services Expense (Acct. 6620)

Using FCC's Regression

	FCC Regression Forecast	Confidence Interval Lower 95%	Upper 95%	Forecast Error %
Expense Assoc. with Switched Access Line	$17.28	$6.02	$28.54	+/- 65%
Expense Assoc. with Special Access Line	$14.06	-$15.38	$43.51	+/- 209%
Weighted Average Access Line Expense	$16.86	$7.26	$26.47	+/- 57%
Expense Assoc. with Toll	$28.26	$18.69	$37.82	+/- 34%
Total Expense	$45.12	$42.75	$47.49	+/- 5%

To provide an alternative more intuitive demonstration of the instability of the FCC's regression results, we created an experiment. In this experiment we used a random process to split the 91 observations used by the FCC into two data sets. We repeated this process five times for each expense account and reran the FCC's first regression equation for each sub-set of the data. Wide swings in the estimated coefficients illustrate the instability in the FCC's estimated coefficients. (See Table 12.) Again, we show the results for the "service expense" (account 6620) to illustrate the relationships in the relevant data. Results for the other accounts show a similar pattern. (See Attachments III and IV.)

To assure ourselves that the instability in the regression coefficients was not the result of the split samples being too small, we examined the consistency of means of each of the

[19] In the FCC's proposed regressions, new variables are created by dividing switched lines, special "lines," and toll DEMs by total lines. The correlation between the variables created from switched lines and DEM variables are much less, but the high correlations in the underlying data continue to plague the regressions.

variables used in the split samples. As shown in Table 13, the mean expense per total line, switched access lines per total line, special access lines per total line and toll DEMS per total line are remarkably stable across samples and across runs, considering the instability of regressions coefficients. The average values in the two samples are, therefore very similar, but the coefficients estimated from the two sets of data are very different. This supports the hypothesis that the source of the instability on the individual coefficients is the high degree of collinearity among explanatory variables. With this level of instability in the coefficient estimates it is impossible to have any confidence in the precision of the FCC's division of expenses between supported and non-supported services.

Table 12

Expense Regression Coefficients Using Split Samples (A and B) for Services (Acct. 6620)

	Coefficient of Switched Line *(Full Sample Coeff. = $17.28)*		Coefficient of Special Line *(Full Sample Coeff. = $14.06)*		Coefficient of Toll MOU *(Full Sample Coeff. = $0.0063)*	
Run #	Sample A	Sample B	Sample A	Sample B	Sample A	Sample B
1	$18.72	$15.18	$6.88	$15.49	$0.0066	$0.0063
2	-$1.62	$25.45	-$1.80	$15.06	$0.0109	$0.0044
3	$4.82	$26.83	-$3.95	$25.21	$0.0094	$0.0042
4	$25.09	$15.76	$37.04	-$7.62	$0.0042	$0.0070
5	$18.68	$15.20	$28.49	-$0.15	$0.0057	$0.0071

Table 13

Variable Means for Split Samples (A and B)

for Services (Acct. 6620)

Run #	Expense/Total Lines (Full Sample Mean = $45.22)		Sw. Lines/Total Lines (Full Sample Mean = $0.90)		Sp. Lines/Total Lines (Full Sample Mean = $0.10)		Toll MOU/Total Lines (Full Sample Mean = $4,474)	
	Sample A	Sample B	Sample A	Sample B	Sample A	Sample B	Sample A	Sample B
1	$46.47	$44.00	0.89	0.91	0.11	0.09	4,386	4,560
2	$45.41	$45.03	0.89	0.90	0.11	0.10	4,315	4,629
3	$44.41	$46.01	0.89	0.90	0.11	0.10	4,309	4,635
4	$44.88	$45.55	0.89	0.91	0.11	0.09	4,398	4,548
5	$44.70	$45.73	0.90	0.90	0.10	0.10	4,399	4,548

D. Other data issues

Another difficulty with the expense data stems from the fact that approximately only twenty local exchange companies account for the 91 observations in the sample. The problem is that the parent companies generally assign a significant portion of non-plant specific and customer operations expenses across their operating companies on the basis of an allocation mechanism. As a result, a simple regression on the 91 observations will produce coefficients that reflect a blend of two relationships: 1) the cost causal relationship, and 2) the allocation-based relationship, of which only the former is appropriate to measure. To net out the latter, it is necessary to either model the allocation method explicitly or, more practically, to aggregate the data to the parent company level. Of course, aggregation of the data would result in a much smaller, albeit truer, data set (20 observations). Furthermore, the consolidated data set does not remedy the multicollinearity problem.

E. After model adjustments

There are numerous after regression adjustments proposed by the FCC to eliminate costs from the cost allocated to switched lines by the faulty regression analysis. Some of these adjustments are large and at least one adjustment results in a double elimination of costs related to special access and toll services. For marketing, the FCC proposes an after model adjustment that eliminates 95.6 percent of the expense that it associates with switched lines. This reduction is based on a study by ETI that estimates that residential marketing expenses are only 4.4 percent of *all* marketing expenses. This adjustment all but overwhelms the regression results, and it is applied incorrectly. With its regression, the FCC eliminates marketing expenses that it estimates are caused by special access and toll services. It appears that the ETI based reduction of 95.6 percent is meant to eliminate some of these same expenses. To the extent that there is overlap between the marketing expense eliminated by the FCC's regressions and the marketing expenses eliminated by the ETI study, the FCC proposes to remove these expenses twice. This is what is known as double counting, or in this case double eliminating. Another problem with using the ETI result for isolating marketing expense that is relevant for universal service is that the ETI result does not provide for marketing expense to single-line business customers.

IV. Conclusions

For over two years, regulatory proceedings across the country have focused on structures of total service long run incremental cost (TSLRIC) models and values for key model inputs. This process has framed the debate over the estimation and reasonableness of values for key model inputs. Into this process, the FCC is now proposing an alternative methodology for estimating input values for cable costs, placement costs, and several categories of non-plant specific expenses. If carefully considered, properly implemented, and clearly communicated, the FCC's proposed regression methodology has the potential of making positive contributions to the debate over the values for key model inputs. Unfortunately, the FCC's analysis is not carefully considered, properly implemented, or clearly communicated.

The cable cost regressions are based on previous analysis by Gabel and Kennedy, but the FCC altered the Gabel and Kennedy analysis to make the model specifications inconsistent and the implementation of the results incorrect. The FCC also abandoned the mainstream OLS regression technique for a technique that has consistently reduced the cost estimates. This downward adjustment is most pronounced in the estimation of buried placement costs. Given the effects of this change in methodology, it is incumbent upon the FCC to carry out a more thorough investigation of the data before resorting to a robust estimator.

Even more egregious is the FCC's fatally flawed methodology for extrapolating placement cost estimates for density zones three through nine. This methodology begins with a cost estimate in density zone two that is below the values from HAI and U S WEST proposed input for the BCPM, is based on nationwide BCPM default values that were disavowed by supporters of this model, and uses information that is, by definition, incorrect. The debate over placement costs, formed over the past two years, requires careful consideration of the facts, not an ad hoc methodology that sidesteps the difficult issues. By sidestepping the difficult issues the FCC's proposed methodology and results detract from the meaningful debate over values for this important model input.

The same can be said for the FCC's proposed regression methodology for assigning expenses to supported services. It is very difficult to assign non-plant specific expenses across services. Rather than further the progress of the debate over the most reasonable assignment of expenses, the FCC's proposed methodology sidesteps the difficult questions with a poorly considered and improperly applied methodology. A fatal flaw in the FCC's methodology is that the underlying data for its explanatory variables are so highly collinear that it is not possible to estimate the causal relationships between specific services and expenses. With no confidence in the precision of the estimated coefficients, there can be no confidence is the accuracy of the assignment of expenses to supported services. Without confidence in the assignment, there is no value in the FCC's proposed methodology. As with the other regression analysis proposed by the FCC, the methodology proposed for assigning non-plant specific expenses detract from the search for realistic values for key model inputs.

Attachment I

GREGORY M. ATTIYEH

Office:
Law & Economics Consulting Group
2000 Powell Street, Suite 600
Emeryville, CA 94608
(510) 450 - 6717

email: greg_attiyeh@lecg.com

Home:
7 Kulani Lane
Pleasant Hill, CA 94523
(925) 280 - 8472

PROFESSIONAL EXPERIENCE

Senior Economist — ***July 1996 - Present***
Law & Economics Consulting Group
Emeryville, CA

- *Financial modeling and valuation for telecommunications industry clients.* Assess financial and competitive impacts of 1) Telecommunications Act of 1996 and FCC's order to open local exchange markets to competition; 2) regulatory delay of Bell entry into long distance; 3) telecommunications deregulation in Canada. Perform business valuations of entry into wireless (PCS), local exchange, and high capacity service markets.
- *Litigation and Strategic support.* Manage case work and preparation of economic testimony for various telecommunications proceedings, including 1) competitive analyses and public interest assessments of Regional Bell filings for long distance entry; 2) rate proceedings; 3) interconnection arbitration proceedings and cost dockets. Wireless spectrum auction tracking and bidding strategy support.
- *Cost modeling.* Develop, document and critique entry models and clients' cost models. Analyze cost trends.
- *Econometric analysis.* Prepared 1) survival analysis and customer churn model to assess viability of new interactive service; 2) time series demand forecasts of Federal Reserve currency; 3) model of industry price trends to measure damages resulting from alleged price fixing.

Consultant, Graduate Record Examination Board — ***June 1995 – June 1998***
Educational Testing Service,
Princeton, NJ

- *Econometric study:* Logistic and survival analysis, using recently collected data from Association of Graduate Schools, to determine factors important in graduate school admissions. Issues include 1) whether factors other than aptitude/achievement, such as gender, ethnicity, and country of citizenship enter into the admissions decision; 2) impact of affirmative action on demand for underrepresented minorities. Published in *Journal of Human Resources*.
- *Follow up study:* Addressed the factors which affect completion of doctoral degree. Issues include extent to which 1) information available at time of admissions forecasts persistence in graduate study; 2) financial assistance influences performance and retention. Forthcoming publication in GRE research journal.

Analyst, Market Analysis and Economics ***June 1995 - June 1996***
AT&T Bell Labs, Business Operations Analysis
Somerset, NJ

- *Computer simulation.* Developed tool to assist regional marketing managers in identifying profitable offer strategies.
- *Quantitative evaluation/forecasting* for acquisition strategy and planning: Identified and countered acquisition strategies of small but fast growing *third tier* long distance competitors.
- *Vulnerability modeling*: Worked with team of analysts to identify the impact of deregulation and entry of resellers into the long distance market in Australia. Designed econometric model to identify customers vulnerable to competitors. Authored internal survey/guide to consumer choice modeling. Internal vulnerability analysis includes modeling to identify AT&T customers likely to switch carriers.
- *Retention analysis:* Projected customer response to promotions.

EDUCATION

Ph.D. , Economics ***1990-1995***
University of Arizona ***Completed in May, 1997***
Tucson, AZ

Dissertation:

- A Theoretical and Experimental Investigation of Volatility Persistence in Financial Markets.
- Fields: Microeconomics, Game Theory, Econometrics. Principal advisor: Vernon Smith.

Teaching:

- Instructor, Intermediate Microeconomics for Business Majors, four semesters.
- Teaching assistant for graduate & undergraduate courses in econometrics, economic principles, intermediate microeconomics.
- Received University of Arizona Foundation's *Outstanding Teaching Award*, 1993.

Research:

- Discrete choice analysis for "Variables Impacting Supply of Minority, Female, and Male Scientists and Engineers." Explained gender and ethnic differences in choice of college major and profession.
- Co-authored "Experiments with the Pivot Process for Providing Public Goods." Tested the practical application of an alternative to contingent valuation methods for eliciting truthful valuations of public goods, for example, to estimate environmental damages. Forthcoming in *Public Choice.*

M.S., Engineering Economics Systems ***June 1989***
Stanford University
Stanford, CA

- Fields: Decision Analysis, Economics.

B.A., Economics ***June 1988***
University of California, Los Angeles
Los Angeles, CA

- Minor in Math and Computer Science.

WILLIAM L. FITZSIMMONS

LECG, Inc.
2000 Powell Street, Suite 600
Emeryville, CA 94608
Tel. (510) 653-9800
Fax (510) 653-9898
E-mail: wlfitz@lecg.com

EDUCATION

Ph.D., Resource Economics, UNIVERSITY OF MASSACHUSETTS, Amherst, MA, 1986

Emphasis: econometrics, natural resource economics, microeconomics, project evaluation, and industrial organization

M.S., Resource Economics, UNIVERSITY OF MASSACHUSETTS, Amherst, MA, 1981

Emphasis: project evaluation, and economics of forestry

B.S., Economics, STATE UNIVERSITY OF NEW YORK AT STONY BROOK, NY, 1975

PRESENT POSITION

LECG, Emeryville, CA, December 1993 - present
Principal, January 1998 - present
Senior Managing Economist, January 1997 - December 1997
Managing Economist, December 1993 - December 1996

- Construct financial simulation models for the analysis of telecommunications issues, including interconnection policies and competitive entry into the local exchange
- Analyze domestic and international telecommunications issues and provide expert witness testimony for regulatory proceedings and litigation
- Work with telecommunications clients to develop and improve cost models
- Assess impacts to telecommunications firms and competition from uneconomic or unlawful policies and practices

PROFESSIONAL EXPERIENCE

BELLSOUTH CORPORATION, Atlanta, GA, January 1988 - December 1993
Senior Economist, April 1992 - December 1993
Corporate Economist, January 1988 - April 1992

- Applied the tools of economic, financial and quantitative analysis to the identification and solution of a broad range of business problems, and developed recommendations for use by senior management in making policy decisions
- Key role in building model of the telephone company that interconnects behavioral equations for capital spending, expenses, real revenues, regulation, and a production function
- Based on model output, formulated and presented policy recommendations and contingency plans to meet expected changes in BellSouth's business environment, such as more severe competition, alternative regulation, and investment in multimedia
- Assessment of potential impacts of wireless on traditional wireline and cellular services
- Analyzed corporate level impacts of prospective mergers and acquisitions
- Derived econometric model that is used to create capital spending targets for the Telco and explore network investment options
- Analyzed corporation's advertising and publishing business to assist with derivation of a new pricing strategy
- Estimated the financial impacts of proposed permutations of interstate price caps
- Provided financial modeling analysis for the tender and bid process for international investments

AT&T, Bedminster, New Jersey, June 1986 - January 1988
Market Analysis and Forecasting

- Developed econometric forecasting models for telecommunication services; identified direction and financial implications of customer migration among private line services; wrote principal components regression software; presented technical and theoretical papers and seminars

PAPERS FILED WITH REGULATORY AGENCIES

Paper prepared for Telecom New Zealand titled "Review of Network Costing Model Used in Todd Telecommunications Consortium Report," by George Barker, William L. Fitzsimmons, Kieran Murray & Graham Scott dated December 2, 1998

"LECG Financial Simulation Model of Effects of FCC Policies on Large Local Exchange Carriers," by Dr. William Fitzsimmons, Dr. Robert Crandall, Professor Robert G. Harris, and Professor Leonard Waverman, Paper filed with FCC, August 1996

PRESENTATIONS AND REGULATORY PROCEEDINGS

Joint reply affidavit with Debra Aron and Robert G. Harris on behalf of Ameritech in the matter of implementation of the local competition provisions in the Telecommunications Act of 1996 (CC Docket No. 96-98); filed June 10, 1999

Expert affidavit on behalf of Ameritech in the matter of implementation of the local competition provisions in the Telecommunications Act of 1996 (CC Docket No. 96-98); filed May 26, 1999

Expert written testimony and cross-examination on behalf of US West in interconnection arbitration proceedings in 1997

South Dakota (Docket No. TC96-184),

Montana (Docket No. D96.11.200),

Wyoming (Docket Nos. 72000-TS-96-95 and 70000-TS-96-319),

New Mexico (Docket No. 96-411-TC),

North Dakota (Docket No. PU-453-96-497),

Idaho (Docket Nos. USW-T-96-15 and ATT-T-96-2), and

Colorado (Docket No. 96S-331T)

Participated in cost workshops on behalf of US West with the Utah Division of Public Utilities and Minnesota Commission in 1996, 1997, and 1998

Expert written testimony and cross-examination on behalf of US West in consolidated cost dockets in

Arizona (Docket Nos. U-3021-96-448, 1996),

Iowa (Docket No. RPU-96-9, 1997),

New Mexico (Docket Nos. 96-310-TC and 97-334-TC, 1998),

Minnesota (Docket Nos. P-442, 5321, 3167, 466, 421/CI-96-1540, 1998), and

Utah (Docket No. 94-999-01, Phase III, Part C, 1998)

Expert testimony and cross-examination in universal service proceedings on behalf of U S WEST in 1997 and 1998

New Mexico (Docket Nos. 96-310-TC, 97-334-TC),

Minnesota (MPUC Docket No. P-999/M-97-909),

Wyoming (General Order No. 81),

Idaho (Case No. GNR-T-97-22), and

Nebraska (Application No. C-1633)

Expert declarations in support of motions for summary judgment by U S WEST in Iowa (June 1997) and Washington (January 1998)

Presentation on "TELRIC Concepts and Applications," Basics of Regulation Conference, New Mexico State University Center for Public Utilities and the National Association of Regulatory Commissioners, Albuquerque, New Mexico, September 18, 1996

June 1999

Attachment II

Table II.1

Predicted Expense and Confidence Intervals for Marketing (Acct. 6610)

	FCC Regression Forecast	Confidence Interval Lower 95%	Upper 95%	Forecast Error %
Expense Assoc. with Switched Access Line	$7.19	$1.24	$13.14	+/- 83%
Expense Assoc. with Special Access Line	$37.64	$22.08	$53.21	+/- 41%
Weighted Average Access Line Expense	$11.10	$6.02	$16.18	+/- 46%
Expense Assoc. with Toll	$8.01	$2.95	$13.06	+/- 63%
Total Expense	$19.10	$17.85	$20.36	+/- 7%

Table II.2

Predicted Expense and Confidence Intervals for Services (Acct. 6620)

	FCC Regression Forecast	Confidence Interval Lower 95%	Upper 95%	Forecast Error %
Expense Assoc. with Switched Access Line	$17.28	$6.02	$28.54	+/- 65%
Expense Assoc. with Special Access Line	$14.06	-$15.38	$43.51	+/- 209%
Weighted Average Access Line Expense	$16.86	$7.26	$26.47	+/- 57%
Expense Assoc. with Toll	$28.26	$18.69	$37.82	+/- 34%
Total Expense	$45.12	$42.75	$47.49	+/- 5%

Table II.3

Predicted Expense and Confidence Intervals

for Corporate Operations (Acct. 6700)

	FCC Regression Forecast	Confidence Interval Lower 95%	Upper 95%	Forecast Error %
Expense Assoc. with Switched Access Line	$34.24	$17.48	$50.99	+/- 49%
Expense Assoc. with Special Access Line	$8.42	-$35.39	$52.24	+/- 520%
Weighted Average Access Line Expense	$30.92	$16.63	$45.22	+/- 46%
Expense Assoc. with Toll	$28.70	$14.47	$42.92	+/- 50%
Total Expense	$59.62	$56.09	$63.14	+/- 6%

Table II.4

Predicted Expense and Confidence Intervals

for Other PP&E (Acct. 6510)

	FCC Regression Forecast	Confidence Interval Lower 95%	Upper 95%	Forecast Error %
Expense Assoc. with Switched Access Line	$0.51	$0.09	$0.94	+/- 83%
Expense Assoc. with Special Access Line	-$0.04	-$1.15	$1.07	+/- 3043%
Weighted Average Access Line Expense	$0.44	$0.08	$0.81	+/- 82%
Expense Assoc. with Toll	-$0.11	-$0.47	$0.25	+/- 328%
Total Expense	$0.33	$0.24	$0.42	+/- 27%

Table II.5

Predicted Expense and Confidence Intervals for Network Operations (Acct. 6530)

	FCC Regression Forecast	Confidence Interval Lower 95%	Upper 95%	Forecast Error %
Expense Assoc. with Switched Access Line	$17.58	$6.78	$28.38	+/- 61%
Expense Assoc. with Special Access Line	$28.82	$0.58	$57.06	+/- 98%
Weighted Average Access Line Expense	$19.02	$9.81	$28.24	+/- 48%
Expense Assoc. with Toll	$15.54	$6.37	$24.71	+/- 59%
Total Expense	$34.57	$32.30	$36.84	+/- 7%

Attachment III

Table III.1

Expense Regression Coefficients Using Split Samples (A and B) for Marketing (Acct. 6610)

Run #	Coefficient of Switched Line (*Full Sample Coeff. = $7.19*)		Coefficient of Special Line (*Full Sample Coeff. = $37.64*)		Coefficient of Toll MOU (*Full Sample Coeff. = $0.0018*)	
	Sample A	Sample B	Sample A	Sample B	Sample A	Sample B
1	-$0.14	$9.92	$33.12	$41.86	$0.0034	$0.0011
2	$4.63	$8.66	$38.19	$35.59	$0.0023	$0.0015
3	$4.74	$10.32	$42.02	$31.33	$0.0021	$0.0014
4	$7.68	$5.94	$21.00	$51.09	$0.0022	$0.0017
5	$7.49	$7.03	$52.40	$22.67	$0.0012	$0.0023

Table III.2

Expense Regression Coefficients Using Split Samples (A and B) for Services (Acct. 6620)

Run #	Coefficient of Switched Line (*Full Sample Coeff. = $17.28*)		Coefficient of Special Line (*Full Sample Coeff. = $14.06*)		Coefficient of Toll MOU (*Full Sample Coeff. = $0.0063*)	
	Sample A	Sample B	Sample A	Sample B	Sample A	Sample B
1	$18.72	$15.18	$6.88	$15.49	$0.0066	$0.0063
2	-$1.62	$25.45	-$1.80	$15.06	$0.0109	$0.0044
3	$4.82	$26.83	-$3.95	$25.21	$0.0094	$0.0042
4	$25.09	$15.76	$37.04	-$7.62	$0.0042	$0.0070
5	$18.68	$15.20	$28.49	-$0.15	$0.0057	$0.0071

Table III.3

Expense Regression Coefficients Using Split Samples (A and B) for Corporate Operations (Acct. 6700)

Run #	Coefficient of Switched Line (*Full Sample Coeff. = $34.24)*) Sample A	Sample B	Coefficient of Special Line (*Full Sample Coeff. = $8.42)*) Sample A	Sample B	Coefficient of Toll MOU (*Full Sample Coeff. = $0.0064)*) Sample A	Sample B
1	$14.25	$41.68	-$7.68	$30.25	$0.0107	$0.0047
2	$19.47	$46.62	$29.58	-$14.11	$0.0086	$0.0048
3	$19.85	$52.64	$30.72	-$23.59	$0.0084	$0.0039
4	$33.14	$36.08	$10.68	$19.82	$0.0059	$0.0065
5	$14.33	$63.87	$5.05	$7.19	$0.0100	$0.0012

Table III.4

Expense Regression Coefficients Using Split Samples (A and B) for Other PP&E (Acct. 6510)

Run #	Coefficient of Switched Line (*Full Sample Coeff. = $0.51)*) Sample A	Sample B	Coefficient of Special Line (*Full Sample Coeff. = -$0.04)*) Sample A	Sample B	Coefficient of Toll MOU (*Full Sample Coeff. = $0.00002)*) Sample A	Sample B
1	$1.17	$0.23	$0.27	-$0.62	-$0.0002	$0.0000
2	$0.86	$0.45	$0.78	-$0.45	-$0.0001	$0.0000
3	$0.95	$0.15	$0.87	-$0.90	-$0.0001	$0.0001
4	$0.70	$0.49	$0.71	-$0.69	-$0.0001	$0.0000
5	$0.95	-$0.12	$0.14	-$0.12	-$0.0001	$0.0001

Table III.5

Expense Regression Coefficients Using Split Samples (A and B) for Network Operations (Acct. 6530)

	Coefficient of Switched Line (*Full Sample Coeff. = $17.58*)		Coefficient of Special Line (*Full Sample Coeff. = $28.82*)		Coefficient of Toll MOU (*Full Sample Coeff. = $0.0035*)	
Run #	Sample A	Sample B	Sample A	Sample B	Sample A	Sample B
1	$25.59	$13.99	$25.31	$39.49	$0.0018	$0.0041
2	$8.09	$19.16	$7.34	$38.43	$0.0064	$0.0026
3	$13.71	$22.06	$24.41	$34.26	$0.0041	$0.0027
4	$34.30	$11.24	$21.25	$34.08	$0.0004	$0.0045
5	$14.38	$22.63	$36.27	$20.40	$0.0037	$0.0029

Attachment IV

Table IV.1
Variable Means for Split Samples (A and B)
for Marketing (Acct. 6610)

	Expense/Total Lines (Full Sample Mean = $18.30)		Sw. Lines/Total Lines (Full Sample Mean = $0.90)		Sp. Lines/Total Lines (Full Sample Mean = $0.10)		Toll MOU/Total Lines (Full Sample Mean = $4,474)	
Run #	Sample A	Sample B	Sample A	Sample B	Sample A	Sample B	Sample A	Sample B
1	$18.53	$18.07	0.89	0.91	0.11	0.09	4,386	4,560
2	$18.28	$18.31	0.89	0.90	0.11	0.10	4,315	4,629
3	$17.92	$18.66	0.89	0.90	0.11	0.10	4,309	4,635
4	$18.97	$17.64	0.89	0.91	0.11	0.09	4,398	4,548
5	$17.41	$19.16	0.90	0.90	0.10	0.10	4,399	4,548

Table IV.2
Variable Means for Split Samples (A and B)
for Services (Acct. 6620)

	Expense/Total Lines (Full Sample Mean = $45.22)		Sw. Lines/Total Lines (Full Sample Mean = $0.90)		Sp. Lines/Total Lines (Full Sample Mean = $0.10)		Toll MOU/Total Lines (Full Sample Mean = $4,474)	
Run #	Sample A	Sample B	Sample A	Sample B	Sample A	Sample B	Sample A	Sample B
1	$46.47	$44.00	0.89	0.91	0.11	0.09	4,386	4,560
2	$45.41	$45.03	0.89	0.90	0.11	0.10	4,315	4,629
3	$44.41	$46.01	0.89	0.90	0.11	0.10	4,309	4,635
4	$44.88	$45.55	0.89	0.91	0.11	0.09	4,398	4,548
5	$44.70	$45.73	0.90	0.90	0.10	0.10	4,399	4,548

Table IV.3
Variable Means for Split Samples (A and B)
for Corporate Operations (Acct. 6700)

	Expense/Total Lines (Full Sample Mean = $60.32)		Sw. Lines/Total Lines (Full Sample Mean = $0.90)		Sp. Lines/Total Lines (Full Sample Mean = $0.10)		Toll MOU/Total Lines (Full Sample Mean = $4,474)	
Run #	Sample A	Sample B	Sample A	Sample B	Sample A	Sample B	Sample A	Sample B
1	$58.65	$61.95	0.89	0.91	0.11	0.09	4,386	4,560
2	$57.61	$62.97	0.89	0.90	0.11	0.10	4,315	4,629
3	$57.31	$63.26	0.89	0.90	0.11	0.10	4,309	4,635
4	$56.46	$64.10	0.89	0.91	0.11	0.09	4,398	4,548
5	$57.31	$63.27	0.90	0.90	0.10	0.10	4,399	4,548

Table IV.4
Variable Means for Split Samples (A and B)
for Other PP&E (Acct. 6510)

	Expense/Total Lines (Full Sample Mean = $0.35)		Sw. Lines/Total Lines (Full Sample Mean = $0.90)		Sp. Lines/Total Lines (Full Sample Mean = $0.10)		Toll MOU/Total Lines (Full Sample Mean = $4,474)	
Run #	Sample A	Sample B	Sample A	Sample B	Sample A	Sample B	Sample A	Sample B
1	$0.41	$0.29	0.89	0.91	0.11	0.09	4,386	4,560
2	$0.30	$0.40	0.89	0.90	0.11	0.10	4,315	4,629
3	$0.40	$0.30	0.89	0.90	0.11	0.10	4,309	4,635
4	$0.34	$0.36	0.89	0.91	0.11	0.09	4,398	4,548
5	$0.39	$0.31	0.90	0.90	0.10	0.10	4,399	4,548

Table IV.5

Variable Means for Split Samples (A and B)

for Network Operations (Acct. 6530)

	Expense/Total Lines (*Full Sample Mean = $34.28*)		Sw. Lines/Total Lines (*Full Sample Mean = $0.90*)		Sp. Lines/Total Lines (*Full Sample Mean = $0.10*)		Toll MOU/Total Lines (*Full Sample Mean = $4,474*)	
Run #	Sample A	Sample B	Sample A	Sample B	Sample A	Sample B	Sample A	Sample B
1	$33.28	$35.24	0.89	0.91	0.11	0.09	4,386	4,560
2	$35.50	$33.08	0.89	0.90	0.11	0.10	4,315	4,629
3	$32.56	$35.96	0.89	0.90	0.11	0.10	4,309	4,635
4	$34.80	$33.76	0.89	0.91	0.11	0.09	4,398	4,548
5	$32.94	$35.58	0.90	0.90	0.10	0.10	4,399	4,548

ATTACHMENT B

Description of SM Clustering Methodology
Using the Denver Curtis Park Wire Center (DNVRCOCP)

Step 0:

The customer geo-coded data is read from the *.in file. The point coordinates are converted from longitude/latitude to relative distances from the central office. Dividing the rectangle that encloses the wire center into squares that are 500 ft. on a side form a starting set of microgrids. The points are then assigned to a specific microgrid. The centroid of the microgrid is calculated using line weighting. If the number of populated microgrids is less than 3000 then the process is complete and the clustering begins. If the number of populated microgrids is greater than 1500 then the length of the microgrid is increased by 500 ft and the process is repeated.

Prior to forming the microgrids, the customer line counts are adjusted to match the entries in the line count table of the database. This includes adjusting the business lines to include public telephone lines and special access channels.

During this process the maximum distance within a microgrid is also calculated. This distance is used to reduce the client input maximum copper distance.

For DNVRCOCP, the statistics were approximately:

14,260 microgrids
760 populated microgrids
500 ft microgrid size
701 ft maximum distance within a microgrid

Step 1:

The microgrids are roughly divided into two groups. Those that are within a copper service area and those that should be served with DLC. The copper customers are in cluster 1 and the DLC customers are within cluster 2. For DNVRCOCP, the effective copper distance limit from the central office for inclusion in cluster 1 is 16,299 feet.

At this stage there are no constraints on line counts or distances, other than the maximum copper distance is reduced by the maximum distance within a microgrid.

Step 2:

At this stage the microgrids in cluster 2 are divided. The new clusters are constrained to meet the distance constraints and also the line limit constraints of a Serving Area Interface (SAI). The line limits are calculated from the adjusted customer line data discussed in Step 0. Default values are 1800 lines with an 80% fill, which yields an effective line constraint of 1440.

These step produces non-contiguous clusters (see clusters 3 through 7 and cluster 9).

Step 3:

This is the 'simple reassignment' phase of the clustering process. In this step, the microgrids are reassigned to the nearest cluster. The cluster centroids are retained from Step 2. There are additional constraints placed on the clustering at this point. For some reason, if a microgrid is to move it must be 1.5 times closer to the new cluster centroid than to the current centroid.

As can be seen from the charts, this step removes the disjointed pockets of customers and non-contiguous clusters left in the Step 2.

Step 4:

This if the 'full optimization' step. If the number of microgrids is less than 1000, then this supposed optimization takes place. It is an attempt to reduce or "tighten up" the clusters. In reality it has little or no effect given the steps that follow.

Step 5:

In this step the clusters with large numbers of lines are split into multiple clusters. The attempt is made to reduce the number of lines to meet the line limit constraint. At this step the client input fill factor is changed to 1.0, effectively increasing the line limit to 1800 lines.

Step 6:

Simple reassignment is applied yet again. See Step 3 above.

Step 7:

The results are written out to the *.clu file. The first portion of this file is a summary of line counts, coordinates and other information by cluster. Before this summary information is written out, each cluster is restored to the original line counts.

Below is a summary of the line counts for each of the clusters for each processing step:

	Processing Step						
Cluster	1	2	3	4	5	6	7
1	84545	84545	66282	66282	16069	16093	7925
2	10756	718	3548	3548	1758	1774	897
3		1439	4527	4527	929	1105	798
4		1438	2845	2845	1045	1026	871
5		1439	4153	4153	600	1308	855
6		1437	3830	3830	288	1552	791

7		1439	2326	2326	529	1379	700
8		1438	3354	3354	1559	1528	772
9		1406	4444	4444	890	1700	955
10					1799	1746	889
11					1796	1617	1169
12					1789	1338	846
13					1799	525	397
14					1735	863	456
15					1779	754	432
16					1635	3287	1766
17					1644	2383	1272
18					1704	685	645
19					1319	1676	1042
20					1534	2550	1336
21					1440	1440	732
22					1032	1032	512
23					1490	1889	944
24					7444	7926	3934
25					2052	2052	1065
26					1002	1091	545
27					1687	1687	833
28					1856	2133	1069
29					3963	3963	1957
30					9704	9704	4780
31					1789	1399	984
32					1799	1674	1478
33					1798	1178	1051
34					1799	1569	886
35					1798	1717	849
36					1764	1770	1011
37					1796	1049	673
38					1735	1607	924
39					1797	852	421
40					1795	1752	899
41					1797	1791	881
42					1756	1119	596

Problems:

1. The cluster process works with the microgrid centroid locations rather than with the customer location directly. The formation of microgrids is not necessary because the clustering process can work directly with customer location data. Apparently, the cluster process applied in this section of the SM is a holdover from some earlier model that was constrained by memory or processing environment.

2. The adjustment to the maximum copper distance is arbitrary. The maximum distance in a given microgrid, point to microgrid centroid, has little or no bearing on the cable distance. The distance constraint for the clustering process needs to be calculated individually for each point in the microgrid.
3. There are problems with the 'simple reassignment' code. If a cluster has to many points and its nearest neighbors are also beyond limits, then nothing is done. At the very least a new cluster should be formed.
4. All of the line adjustments that take place within the model should be consolidated. It is very confusing, apparently even to the SM modelers, how the lines are adjusted. As a matter of fact, the CLUSINTF process should be entirely eliminated. It would make much more sense to consolidate the CLUSTER and CLUSINTF processes in one process.
5. There is no recognition in the SM clustering of the wire center boundaries. Looking at cluster 4 on chart 6 you can see that the distribution plant for this cluster will leave the wire center boundaries. It is also apparent that there is a strong possibility that the feeder plant would escape the wire center boundaries.
6. The division of customers into copper/DLC areas is wasted. When the 'simple reassignment' step, Step 3, is completed any distinction between copper and DLC is gone. To be done properly, the customers served on copper and the customers served with DLC need to be treated separately. This would call for a redesign of the SM clustering process.
7. Changing the fill factor in mid-process is flawed. Whatever factor is used, it should be used consistently. By changing the factor in mid-process, the number of clusters generated is reduced thus reducing the total loop cost.